

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS**

NUANCE COMMUNICATIONS, INC.,

Plaintiff and Counterclaim
Defendant,

v.

OMILIA NATURAL LANGUAGE
SOLUTIONS, LTD.,

Defendant and Counterclaim
Plaintiff.

Case No. 1:19-CV-11438-PBS

OMILIA NATURAL LANGUAGE SOLUTIONS, LTD.’S
NOTICE OF FILING

Omilia Natural Language Solutions, Ltd. (“Omilia”), hereby submits to the Court the slides Omilia used at the Claim Construction Hearing held on July 10, 2020. Omilia submitted these slides to the Court by email on July 9, 2020, at the request of the Court.

Dated: July 15, 2020

Respectfully Submitted,

By /s/ Daniel S. Sternberg

Kevin C. Adam (SBN 684955)
Daniel S. Sternberg (SBN 688842)
WHITE & CASE LLP
75 State Street, 24th Floor
Boston, MA 02109
(617) 979-9300
Kevin.adam@whitecase.com
Daniel.sternberg@whitecase.com

Of Counsel:
Dimitrios Drivas (*admitted pro hac vice*)
Raj Gandesha (*admitted pro hac vice*)
Stefan Mentzer (*admitted pro hac vice*)
John Padro (*admitted pro hac vice*)
WHITE & CASE LLP

1221 Avenue of the Americas
New York, NY 10020-1095
(212) 819-8286
ddrivas@whitecase.com
rgandesha@whitecase.com
smentzer@whitecase.com
john.padro@whitecase.com

Hallie Kiernan (*admitted pro hac vice*)
WHITE & CASE LLP
3000 El Camino Real
Two Palo Alto Square, Suite 900
Palo Alto, CA 94306
(650) 213-0300
hallie.kiernan@whitecase.com

*Counsel for Omilia Natural Language
Solutions, Ltd.*

CERTIFICATE OF SERVICE

I hereby certify that counsel of record who are deemed to have consented to electronic service are being served on July 15, 2020 with a copy of this document via the Court's CM/ECF system per Local Rule CV-5.4 (c).

/s/ Daniel S. Sternberg
Daniel S. Sternberg

Markman Hearing Slides
Nuance Communications, Inc., v.
Omilia Natural Language Solutions, LTD., Case No. 1:19-cv-11438-PBS

Presented by


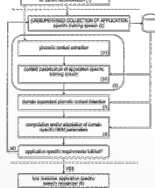
Omilia | Technology that Listens,
Understands and Cares

Introductory Remarks

Presentation by Omilia
No. Case No. 1:19-cv-11438-PBS
United States District Court
District of Massachusetts

Patents Asserted by Nuance


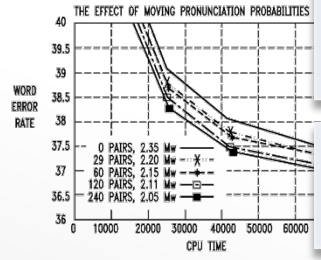
The 6,999,925 Patent

 US00699925B2	
(12) United States Patent Fischer et al.	
(10) Patent No.: (45) Date of Patent:	US 6,999,925 B2 Feb. 14, 2006
(54) METHOD AND APPARATUS FOR PHONETIC CONTEXT ADAPTATION FOR IMPROVED SPEECH RECOGNITION	
(75) Inventors: Volker Fischer, Leimen (DE); Siegfried Kunzmann, Heidelberg (DE); Eric-W. Janke, Winchester (GB); A. Jon Tyrrell, Chancellors Road Eastleigh (GB)	
(73) Assignee: International Business Machines Corporation, Armonk, NY (US)	
(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 698 days.	
(21) Appl. No.: 10/007,990	
(22) Filed: Nov. 13, 2001	
(65) Prior Publication Data US 2002/0087314 A1 Jan. 4, 2002	
(30) Foreign Application Priority Data Nov. 14, 2000 (EP) 00124795	
(51) Int. Cl. <i>G10L 15/06</i> (2006.01) <i>G10L 15/08</i> (2006.01) <i>G10L 15/27</i> (2006.01) <i>G10L 15/18</i> (2006.01)	(57) ABSTRACT The present invention provides a computerized method and apparatus for automatically generating from a first speech recognizer a second speech recognizer which can be adapted to a specific domain. The first speech recognizer can include a first acoustic model with a first decision network and corresponding first phonetic contexts. The first acoustic model can be used as a starting point for the adaptation process. A second acoustic model with a second decision network and corresponding second speech recognizer is generated based on the first decision network and corresponding second speech recognizer.
(52) U.S. CL. 704/243; 704/255; 704/257; 704/236; 704/257, 10, 8 See application file for complete search history.	
(56) References Cited U.S. PATENT DOCUMENTS 5,794,192 A * 8/1998 Zhao 704/244 5,799,277 A * 8/1998 Takami 704/256 6,014,624 A * 1/2000 Raman 704/243	
29 Claims	
	

'925 Patent filed by IBM on Nov. 13, 2001

'925 Patent acquired by Nuance on Dec. 31, 2008

The 8,532,993 Patent

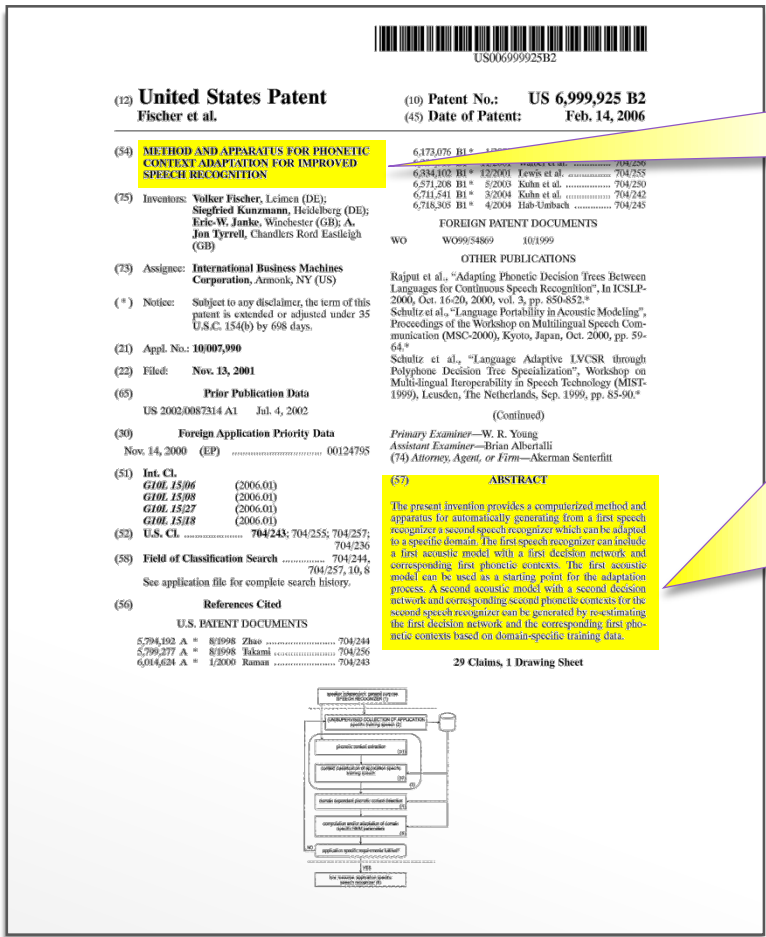
 US008532993B2	
(12) United States Patent Ljolje	
(10) Patent No.: (45) Date of Patent:	US 8,532,993 B2 *Sep. 10, 2013
(54) SPEECH RECOGNITION BASED ON PRONUNCIATION MODELING	
(75) Inventor: Andrej Ljolje, Morristown, NJ (US)	
(73) Assignee: AT&T Intellectual Property II, L.P., Atlanta, GA (US)	
(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days. This patent is subject to a terminal disclaimer.	
(21) Appl. No.: 13/839,996	
(22) Filed: Jul. 2, 2012	
(65) Prior Publication Data US 2012/0271635 A1 Oct. 25, 2012	
Related U.S. Application Data (63) Continuation of application No. 11/380,502, filed on Apr. 27, 2006, now Pat. No. 8,214,213.	
(51) Int. Cl. <i>G10L 15/02</i> (2006.01) <i>G10L 15/04</i> (2006.01) <i>G10L 15/06</i> (2006.01) <i>G10L 15/20</i> (2006.01)	(57) ABSTRACT A system and method for performing speech recognition is disclosed. The method comprises receiving an utterance, applying the utterance to a recognizer with a language model having pronunciation probabilities associated with unique word identifiers for words given their pronunciations and presenting a recognition result for the utterance. Recognition improvement is found by moving a pronunciation model from a dictionary to the language model.
(52) U.S. CL. 704/251; 704/253; 704/254; 704/255; 704/256; 704/257; 704/240	
(58) Field of Classification Search USPC 704/251, 246, 253-257, 240 See application file for complete search history.	
20 Claims	
	

'993 Patent filed by AT&T on Jul. 2, 2012

'993 Patent acquired by Nuance on Dec. 14, 2016

The '925 Patent: Create a New Speech Recognizer By Adapting an Existing One With Updated Phonetic Contexts

The 6,999,925 Patent



METHOD AND APPARATUS FOR **PHONETIC CONTEXT ADAPTATION** FOR IMPROVED SPEECH RECOGNITION

ABSTRACT

The present invention provides a computerized method and apparatus for automatically generating from a first speech recognizer a second speech recognizer which can be adapted to a specific domain. The first speech recognizer can include a first acoustic model with a first decision network and corresponding first phonetic contexts. The first acoustic model can be used as a starting point for the adaptation process. A second acoustic model with a second decision network and corresponding second phonetic contexts for the second speech recognizer can be generated by re-estimating the first decision network and the corresponding first phonetic contexts based on domain-specific training data.

The Specific Method Described In The '925 Patent

1. A computerized method of **automatically generating from a first speech recognizer a second speech recognizer**, said first speech recognizer comprising a first acoustic model with a first decision network and corresponding first phonetic contexts, and **said second speech recognizer being adapted to a specific domain**, said method **comprising**:

based on said first acoustic model, **generating a second acoustic model** with a second decision network and corresponding second phonetic contexts for said second speech recognizer **by re-estimating said first decision network and said corresponding first phonetic contexts** based on domain-specific training data, wherein said first decision network and said second decision network utilize a phonetic decision free to perform speech recognition operations, wherein the number of nodes in the second decision network is not fixed by the number of nodes in the first decision network, and wherein said re-estimating comprises partitioning said training data using said first decision network of said first speech recognizer.

'925 Patent at Claim 1 (emphasis added) (Claim 27 differs)



Sample Decision Network for Bags or Bats

Decision Tree
for “b”



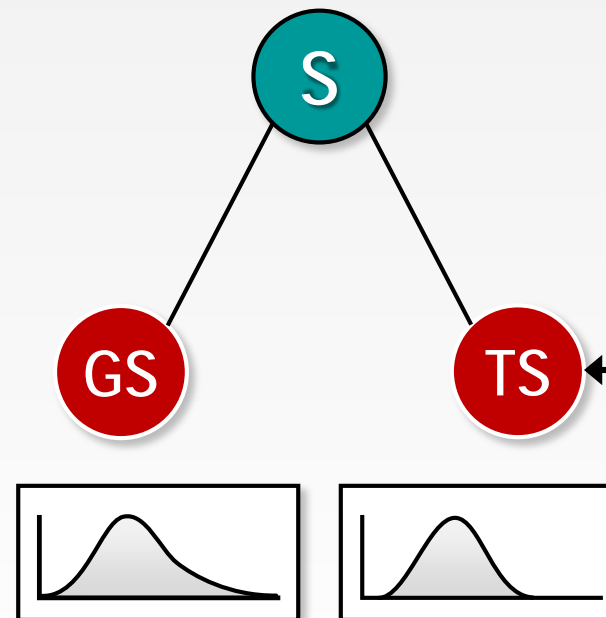
Decision Tree
for “a”



Decision Tree
for “g” or “t”



Decision Tree for “s”



Nodes Tied to Phonetic
Contexts or Question:
For example --
Before “s” is a “g”?
If yes, go left, if no, go right.

Leaf Node (red) –
e.g. Bi-phone

Probability Distribution
or Gaussian

Adaptation of an Acoustic Model Based on New Training Data

- 1. Modify the existing decision network and phonetic contexts**
 - a. Remove phonetic contexts for example, if no longer needed**
 - b. Add new phonetic contexts based on new training data**
- 2. Modify the model (HMM) parameters / probabilities (doesn't change the phonetic contexts)**

For example:

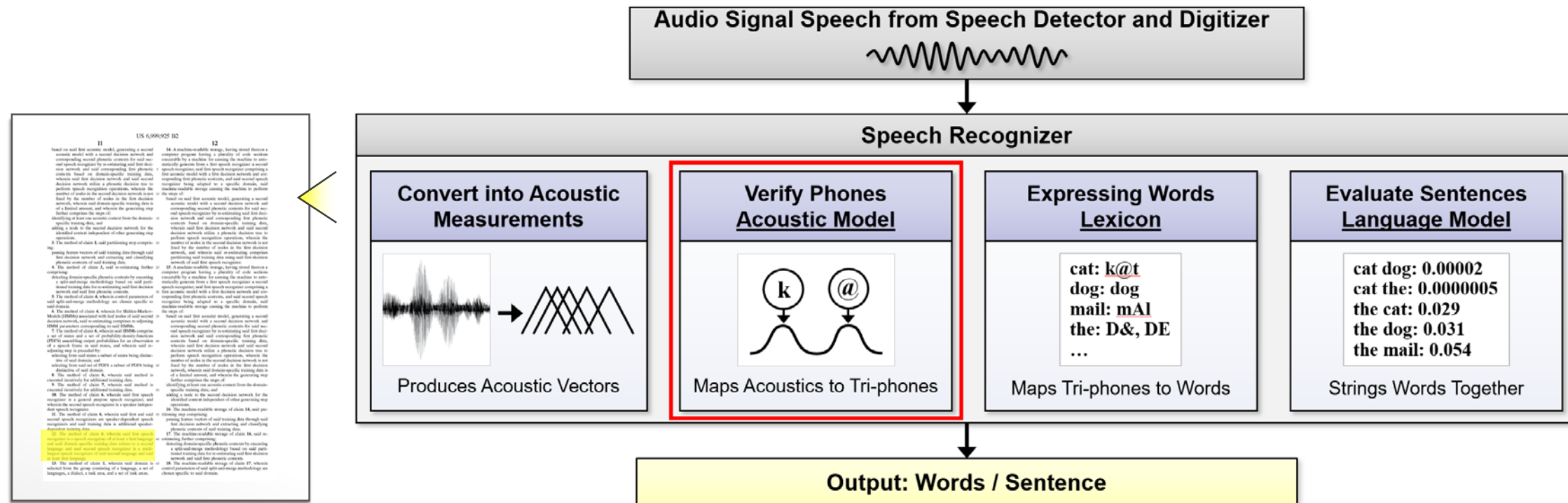
- Take a general English speech recognizer, and use bank specific training data to create an English banking model
 - Might just remove the non-banking phonetic contexts
 - Might need to add phonetic contexts given unique words used in bank setting
- Make a user-specific recognizer. Start with a general English recognizer and use user specific training data. Only update the model probabilities based on the specific user

The Specific Method Described In The '925 Patent

12. The method of claim 6, wherein said first speech recognizer is a speech recognizer of at least a first language and said domain specific training data relates to

③ a second language and said second speech recognizer is a multi-lingual speech recognizer of said second language and said at least first language. ④

'925 Patent at Claim 12 (emphasis added) (similiar elements in Claim 27)



“Second Language:” Proposed Constructions

Claim Term (Purple) in Context	Asserted Claims	Omilia’s Proposed Construction	Nuance’s Proposed Construction
“wherein said first speech recognizer is a speech recognizer of at least a first language and said domain specific training data relates to a second language and said speech recognizer is a multilingual speech recognizer of said second language and said at least first language.”	925: 12, 25	“a second language not present/incorporated in the first speech recognizer”	“a language other than the first language of the first domain/speech recognizer”

“Second Language:” Proposed Constructions

Claim Term (Purple) in Context	Asserted Claims	Omilia’s Proposed Construction	Nuance’s Proposed Construction
“wherein the first domain comprises at least a first language, wherein the second domain comprises at least a second language , and wherein the second speech recognizer is a multi-lingual speech recognizer”	925: 27	“a second language not present/incorporated in the first speech recognizer”	“a language other than the first language of the first domain/speech recognizer”

“Second Language” Means

A Second Language Not Present/Incorporated in the First Speech Recognizer

wherein said first speech recognizer is a speech recognizer of at least a first language and said domain specific training data relates to a **second language** and said speech recognizer is a multilingual speech recognizer of said **second language** and said at least first language

'925 patent at 11:59-64 (emphasis added)

The dispute here is whether the second speech recognizer must be able to recognize something the first recognizer cannot

Omilia: Claims and specification make clear that the second speech recognizer must be able to recognize a new additional language not present in the first recognizer

Nuance: Second speech recognizer need only recognize a language other than the “first” language of the first speech recognizer *even if that other language is present in and recognized by the first recognizer*

“Second Language” Means

A Second Language Not Present/Incorporated in the First Speech Recognizer

Omilia’s proposed construction is supported by:

- 1. The claim language**
- 2. The specification**
- 3. The prosecution history**

Languages May Share Some Common Aspects

Step 2: Detect Phonemes in the Audio File [Acoustic Model]

- A. Each language consists of a fixed number of sounds, called phonemes. English has about 45 phonemes which don't correspond to the 26 letters of the alphabet. Any word in a language can only be composed by combining the phonemes in that language only.
- B. The acoustic realization of a phoneme depends strongly on the context in which it occurs.
 - And different people pronounce things (phonemes) differently, and a speech recognizer can be flexible enough to take this into account
 - Hence, training data, in the form of actual recordings of different people speaking, is necessary to capture the range of pronunciations.
- C. Languages do not recognize the same phonemes but many are overlapping.
 - The phonemes /th/ as in "the" is a sound in English that is not used in many other languages. Non-native speakers may pronounce "this" as "zis" or "dis." The /th/ phoneme is not in French or German.
 - English does not distinguish between two /n/ sounds in Spanish.

A Language May Contain Loan Words

taxi

ok

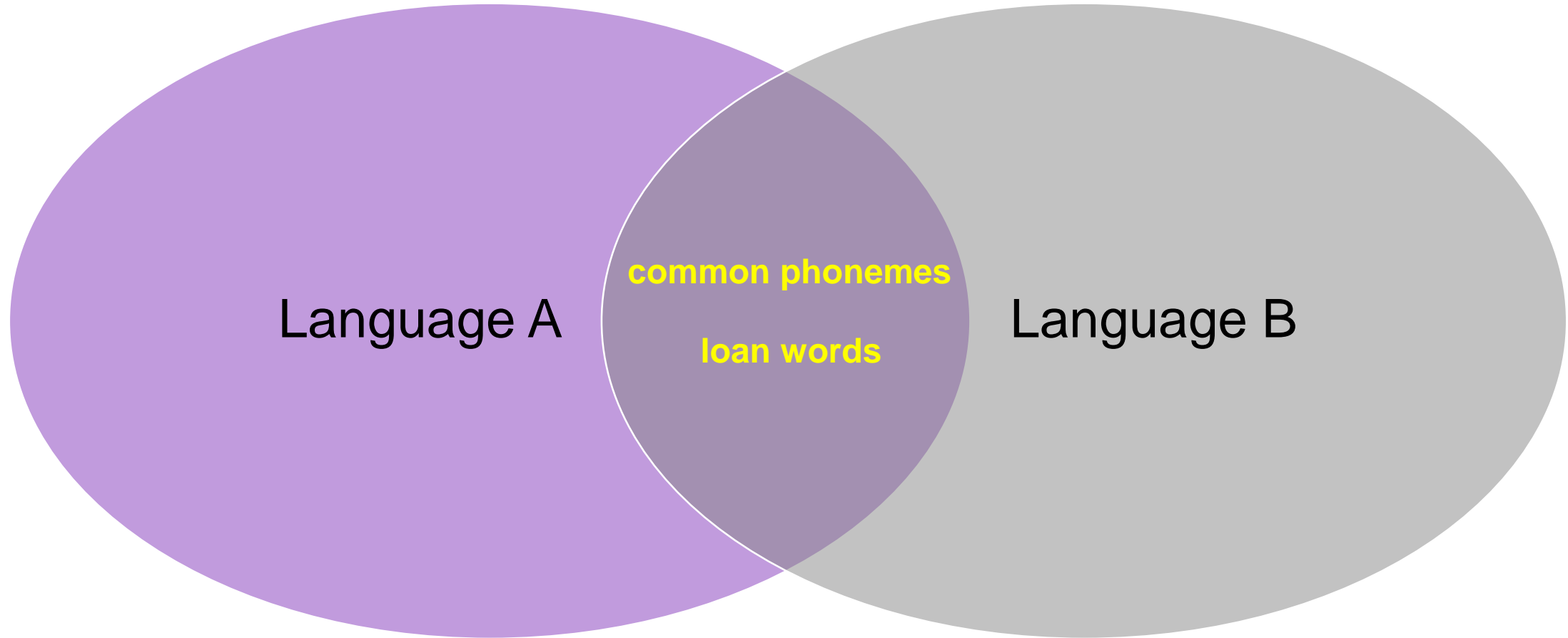
bye bye

tiramisu

Loan Words are Subject to the Phonetic Context of the Language

jalapeño v. jalapeno

Languages Share Some But Not All Aspects



Omilia's Construction is Necessary To Give Meaning to the Claims

Claim 27

A computerized method of generating a second speech recognizer comprising the steps of:

Identifying a first speech recognizer of a first domain comprising a first acoustic model with a first decision network and corresponding first phonetic contexts;

Receiving domain-specific training data of a second domain; and

Based on the first speech recognizer and the domain-specific training data, generating a second acoustic model of said first domain and said second domain comprising a second acoustic model with a second decision network and corresponding second phonetic contexts, **wherein the first domain comprises at least a first language**, wherein **the second domain comprises at least a second language**, and wherein the second speech recognizer is a multi-lingual speech recognizer.

'925 patent at 14:8-23 (emphasis added)

Defines the **first domain as comprising at least a **first language**.**

Distinguishes the **second domain as comprising at least a **second language**.**

First domain does not include the second language.

Second speech recognizer encompasses acoustic models for **first and second domains.**

Omilia's Construction is Necessary To Give Meaning to the Claims

Claim 12 (representative of claim 25)

The method of claim 6 [of automatically generating from a first speech recognizer a second speech recognizer], wherein said **first speech recognizer is a speech recognizer of at least a first language** and said **domain specific training data relates to a second language** and said second speech recognizer is a multi-lingual speech recognizer **of said second language and said at least first language**.

'925 patent at 11:59-64 (emphasis added)

Second domain is a **second language** not included in the first domain and added to the speech recognizer.

Specification Makes Clear That Second Domain (Language) Must be a New Additional Domain (Language) Not Present in the First

Citation	Statement
2:47-56	“In contrast to state of the art approaches, the present invention simultaneously allows for the creation of new phonetic contexts that need not be present in the original training material . Thus, rather than create a domain specific inventory from scratch according to the state of the art, which would require the collection of a huge amount of domain-specific training data, according to the present invention, the inventory of the general recognizer can be adapted to a new domain based on a small amount of adaptation data. ”
6:11-20	“The invention disclosed herein can utilize the already existing phonetic context inventory of a (general purpose) speech recognizer and some small amount of domain specific adaptation data for both the emphasis of dominant contexts and the creation of new phonetic contexts that are relevant for a given domain. ”
8:23-30	“[T]he method of the present invention simultaneously allows the creation of new phonetic contexts that need not be present in the original training material . Rather than create a domain specific HMM inventory from scratch according to the state of the art, which requires the collection of a huge amount of domain-specific training data, the present invention allows the adaptation of the general recognizer’s HMM inventory to a new domain based on a small amount of adaptation data.”
9:19-24	“If different languages share a common phonetic alphabet, the method also can be used for the incremental and data driven incorporation of a new language into a true multi-lingual speech recognizer ”

The Prosecution History Supports Omilia's Construction

“Waibel et al. discloses the creation of a second speech recognizer from a first recognizer [and] adapts the second speech recognizer for a ‘**new, smaller domain**’ (column 6, lines 37-39). Therefore, creating from a first speech recognizer of a first language, with training data from a second language, a second recognizer that is able to recognize **at least the first language and the second language** is not disclosed in the prior art of record and would not have been obvious to one of ordinary skill in the art at the time of the invention.”

‘925 File History, Nov. 3, 2004 Office Action, at 8 (emphasis added)

“[T]he second speech recognizer generated from the first recognizer and the domain-specific data has **a larger domain than the first speech recognizer.**”

‘925 File History, Feb. 3, 2005 Applicant Response (emphasis added).

Nuance Admits The Second Speech Recognizer Must Be Different From the First

“Omilia argues that, ‘[i]f the first speech recognizer already recognized the second language, it would be no different from the newly-generated speech recognizer.’ **Nuance does not necessarily disagree that the resulting speech recognizer must be different from the first**”

Nuance’s Responsive Brief at 11 (citation omitted) (emphasis added)

Nuance’s construction: second recognizer could recognize everything recognized by the first ***and only that***.

Omilia’s construction: captures the fact that the second recognizer is different from the first because it can recognize a second language that the first recognizer could not.

“Multi-Lingual Speech Recognizer:” Proposed Constructions

Claim Term (Purple) in Context	Asserted Claims	Omilia’s Proposed Construction	Nuance’s Proposed Construction
“wherein said first speech recognizer is a speech recognizer of at least a first language and said domain specific training data relates to a second language and said speech recognizer is a multilingual speech recognizer of said second language and said at least first language.”	925: 12, 25	“a speech recognizer incorporating at least the first language and adding the second language (i.e., not present/incorporated in the first speech recognizer)”	“a speech recognizer with an acoustic model that covers sound units of multiple languages”

“Multi-Lingual Speech Recognizer:” Proposed Constructions

Claim Term (Purple) in Context	Asserted Claims	Omilia’s Proposed Construction	Nuance’s Proposed Construction
“wherein the first domain comprises at least a first language, wherein the second domain comprises at least a second language, and wherein the second speech recognizer is a multi-lingual speech recognizer ”	925: 27	“a speech recognizer incorporating at least the first language and adding the second language (i.e., not present/incorporated in the first speech recognizer)”	“a speech recognizer with an acoustic model that covers sound units of multiple languages”

A “Multi-Lingual Speech Recognizer” Must Recognize a Second Language Not Present/Incorporated in the First Speech Recognizer

wherein the first domain comprises at least a first language, wherein the second domain comprises at least a second language, and wherein the second speech recognizer is **a multi-lingual speech recognizer**”

’925 patent claim at 11:59-64 (emphasis added)

The dispute: whether a speech recognizer must recognize more than just some sound units common to a first and second language to be a “multi-lingual speech recognizer”

Omilia: Claims and specification make clear that a multi-lingual speech recognizer must be able to recognize a new additional language not present in the first recognizer

Nuance: A speech recognizer need only recognize “sound units” common to two languages to be a “multi-lingual speech recognizer”

Omilia's Construction is Necessary To Give Meaning to the Claims

Claim 27

A computerized method of generating a second speech recognizer comprising the steps of:

Identifying a first speech recognizer of a first domain comprising a first acoustic model with a first decision network and corresponding first phonetic contexts;

Receiving domain-specific training data of a second domain; and

Based on the first speech recognizer and the domain-specific training data, generating a second acoustic model of said first domain and said second domain comprising a second acoustic model with a second decision network and corresponding second phonetic contexts, **wherein the first domain comprises at least a first language, wherein the second domain comprises at least a second language, and wherein the second speech recognizer is a multi-lingual speech recognizer**

'925 patent at 14:8-23 (emphasis added)

Claim 12 (representative of claim 25)

The method of claim 6 [of automatically generating from a first speech recognizer a second speech recognizer], wherein said first speech recognizer is a speech recognizer of at least a first language and said domain specific training data relates to a second language **and said second speech recognizer is a multi-lingual speech recognizer of said second language and said at least first language.**

'925 patent at 11:59-64 (emphasis added)

Case 1:19-cv-11438-PBS Document 1-1 Filed 07/15/20 Page 28 of 51

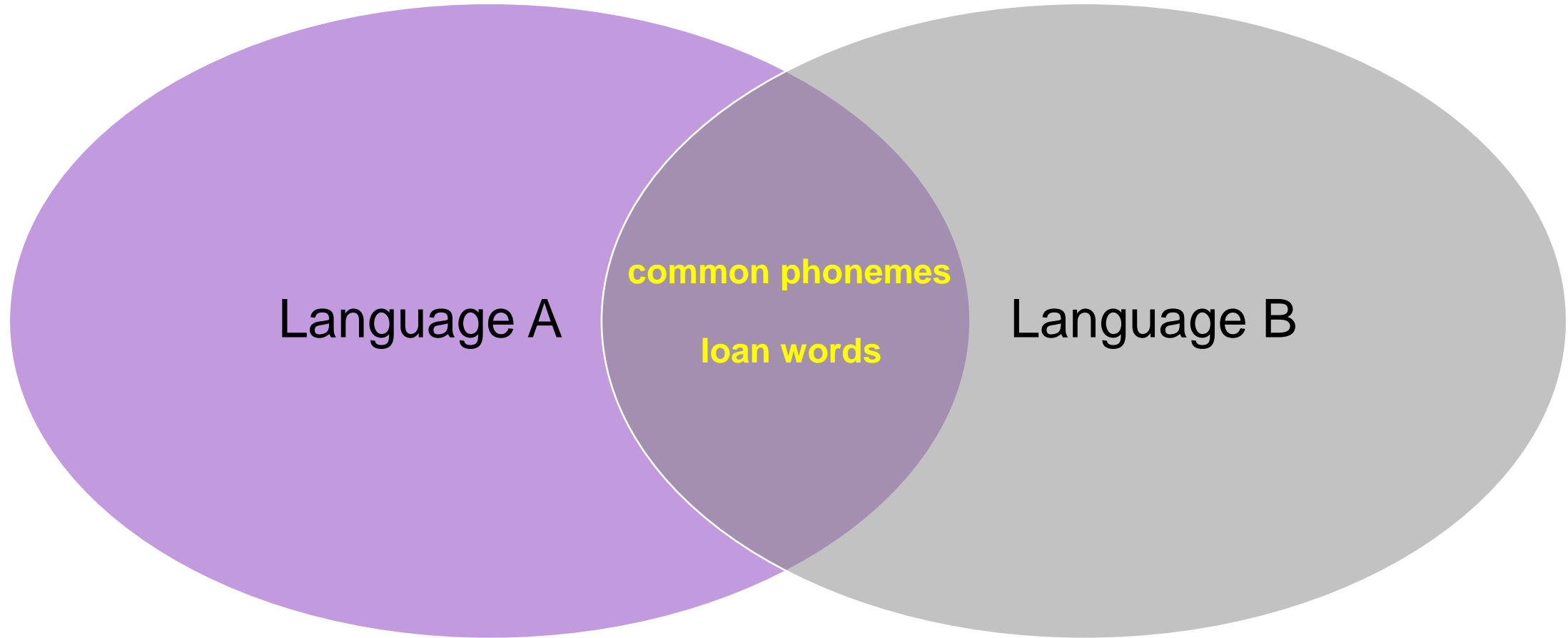
Specification Makes Clear That A Multi-Lingual Speech Recognize Can Recognize a First Language and a New Additional Second Language

“If different languages share a common phonetic alphabet, the method also can be used for the incremental and data driven incorporation of **a *new language* into a true multi-lingual speech recognizer** that shares HMMs between languages.”

‘925 patent 9:19-23 (emphasis added)

By focusing on “sound units,” Nuance’s proposed construction reads “language” out of “multi-lingual.”

Languages Share Some But Not All Aspects



Nuance's Extrinsic Evidence

“[A] multi-lingual speech recognizer for English and French would include some phonemes (sound units) that are shared between the two languages, like /a/, /u/, and /f/, but **would typically also include sound units that are used in one language but not the other.** In the case of English and French, for example, the recognizer would include the English phoneme /θ/ (the first consonant sound in “thief”), which is not used in French, as well as the French phoneme /œ/ (the initial vowel in “oeuf”), which is not used in English.”

Declaration of Karen Livescu ¶ 46 (emphasis added)

“Generating a Second Acoustic Model”

Presentation by Omilia
No. Case No. 1:19-cv-11438-PBS
United States District Court
District of Massachusetts

Proposed Constructions

Claim Term (Purple) in Context	Asserted Claims	Omilia's Proposed Construction	Nuance's Proposed Construction
<p>“based on the first speech recognizer and the domain-specific training data, generating a second acoustic model of said first domain and said second domain comprising a second acoustic model with a second decision network and corresponding second phonetic contexts”</p>	925: 27	<p>“generating a second acoustic model” requires “re-estimating said first decision network and said corresponding phonetic contexts”</p>	plain and ordinary meaning

Case 1:19-cv-01438-PBS Document 141 Filed 07/25/20 Page 33 of 52

“Generating a Second Acoustic Model” Requires Re-estimation of Phonetic Contexts

based on the first speech recognizer and the domain-specific training data, **generating a second acoustic model** of said first domain and said second domain comprising a second acoustic model with a second decision network and corresponding second phonetic contexts

'925 patent at 14:15-20 (emphasis added)

What does claim 27 mean by “generating a second acoustic model”?

Omilia: Generating by “re-estimating said first decision network and said corresponding phonetic contexts” as described by the inventors as their invention and distinct from other modes of generating acoustic models

OR

Nuance: No definition, i.e. plain meaning

Case 1:19-cv-01438-PBS Document 1-1 Filed 07/16/20 Page 34 of 52

Nuance's Plain Meaning Construction for “Generating a Second Acoustic Model” Is Incorrect

1. **“Plain” meaning renders the term meaningless; no definition**
2. **Potentially covers any “generating” of a second acoustic model, independent of whether the applicants here invented it:**
 - A. **Contrary to the specification and the inventors’ representations that other generation forms are not their invention and what is their invention**
 - B. **Renders claims invalid – no description or enablement supporting it**
 - C. **Based on extrinsic evidence, not the intrinsic record in violation of the law**
3. **Violates the core quid pro quo of patents – you get a patent for what it is you disclose; cannot change that during litigation**

Case 1:19-cv-01438-PBS Document 141 Filed 07/15/20 Page 33 of 52

“Generating a Second Acoustic Model” Requires Re-estimation of Phonetic Contexts

Law is clear that Omilia’s construction is correct -- generating by “re-estimating said first decision network and said corresponding phonetic contexts”:

- 1. Specification defines “generation” by re-estimation of first phonetic contexts as the “invention”**
- 2. Specification distinguishes this type of “generating acoustic models” from other known generating techniques in the prior art**
- 3. Context of claim and prosecution confirm claim 27 concerns “generation” by re-estimation of the first phonetic contexts**

Omilia's Construction is Necessary To Give Meaning to the Claim

Parts of Claim 27

identifying a first speech recognizer of a **first domain** comprising a first acoustic model with a first decision network and corresponding first phonetic contexts;
receiving domain-specific training data of a second domain; and

- ① **based on the first speech recognizer** and the domain-specific training data, **generating a second acoustic model** of said **first domain and said second domain** comprising a second acoustic model with a second decision network and corresponding second phonetic contexts, wherein the **first domain comprises at least a first language, wherein the second domain comprises at least a second language**, and wherein the second speech recognizer is a multi-lingual speech recognizer.
- ②

'925 patent at 14:15-23 (emphasis added)

- ① Without “by re-estimating said first decision network and said corresponding first phonetic context,” claim provides no explanation as to what it means to be “based on the first speech recognizer”
- ② Without “re-estimating...phonetic contexts” no way to add a “second language” – additional language requires new language-specific phonetic contexts

Case 1:19-cv-01138-PBS Document 141 Filed 07/15/20 Page 37 of 52

Specification Defines the Invention as Adapting the Phonetic Context

Orthogonally to these previous approaches, the **present invention focuses** on the re-estimation of phonetic contexts, or-in other words-the adaptation of the recognizer's sub-word inventory to a special domain. Whereas in any speaker adaptation algorithm, as well as in the above mentioned documents of V. Fischer et al., the phonetic contexts once estimated by the training procedure are fixed, the present invention utilizes a small amount of upfront training data for the domain specific insertion, deletion, or adaptation of phones in their respective context. Thus re-estimation of the phonetic contexts refers to a (complete) recalculation of the decision network and its corresponding phonetic contexts based on the general speech recognizer decision network.

'925 patent at 6:66-7:12 (emphasis added)

The invention disclosed herein can utilize the already existing phonetic context inventory of a (general purpose) speech recognizer and some small amount of domain specific adaptation data for both the emphasis of dominant contexts and the creation of new phonetic contexts that are relevant for a given domain. **This is achieved by** using the speech recognizer's decision network and its corresponding phonetic contexts as a starting point and by re-estimating the decision network and phonetic contexts based on domain-specific training data.

'925 patent at 6:11-20 (emphasis added)

- **See also 7:37-51, 8:14-30; Title** (“Method and Apparatus for **phonetic context adaptation** for improved speech recognition”)
- **Re-estimation of the first phonetic contexts also described as the invention in the:**
 - **Summary of the invention** (2:31-56)
 - **Abstract**

Case 1:19-cv-01138-PSS Document 141 Filed 07/15/20 Page 39 of 52

Specification Distinguished its Generation from Different Prior Art Generation of Acoustic Models

- **Repeatedly distinguishes its adapting of the first phonetic contexts from other ways to generate or adapt acoustic models:**

Citation	Definitional Statement
2:31-56	summary of the invention; the method of preserving and adding new phonetic contexts “in contrast to the prior art”
3:41-5:26	describing the known methods for generating speech recognizers “from scratch” including “bootstrapping”; distinguished in summary of invention (2:31-56; also 8:14-30) from the inventions’ adaptation of phonetic contexts
6:32-46	admitting that adaptation of acoustic model (HMM) parameters were known and distinguishing that prior art from changing the “phonetic contexts” (i.e. what the patent describes as the invention)
6:66-7:17	distinguishing the “adaptation” of the patent from what was known because it is a re-estimation of the decision network and phonetic contexts
8:14-30	Again distinguishing its adaptation from the known manners to generate new acoustic models because it preserves the original phonetic contexts and can add phonetic contexts (referencing 5:29-66)

- **Exactly why the term must be construed as Omilia proposes – cannot now claim an interpretation at odds with what is distinguished in the specification**
- **Nuance presented no response to this in their briefing**

No Other “Generating” Described as Part of the Invention

- **Generation by re-estimation, starting with existing phonetic context and adapting or re-estimating them, is what provides the alleged benefits of the patent: save time and money**
2:44-56; 6:11-31; 8:14-40

- **Only type of “generation” described in specification as its invention is “by re-estimation of the first decision network and corresponding phonetic contexts”**

Fig. 1, step 4; 7:62-8:10 (describing the re-estimation process)

- **Other manners of adaptation discussed and distinguished from the present invention**

- Adjusting HMM parameters (6:31-45)

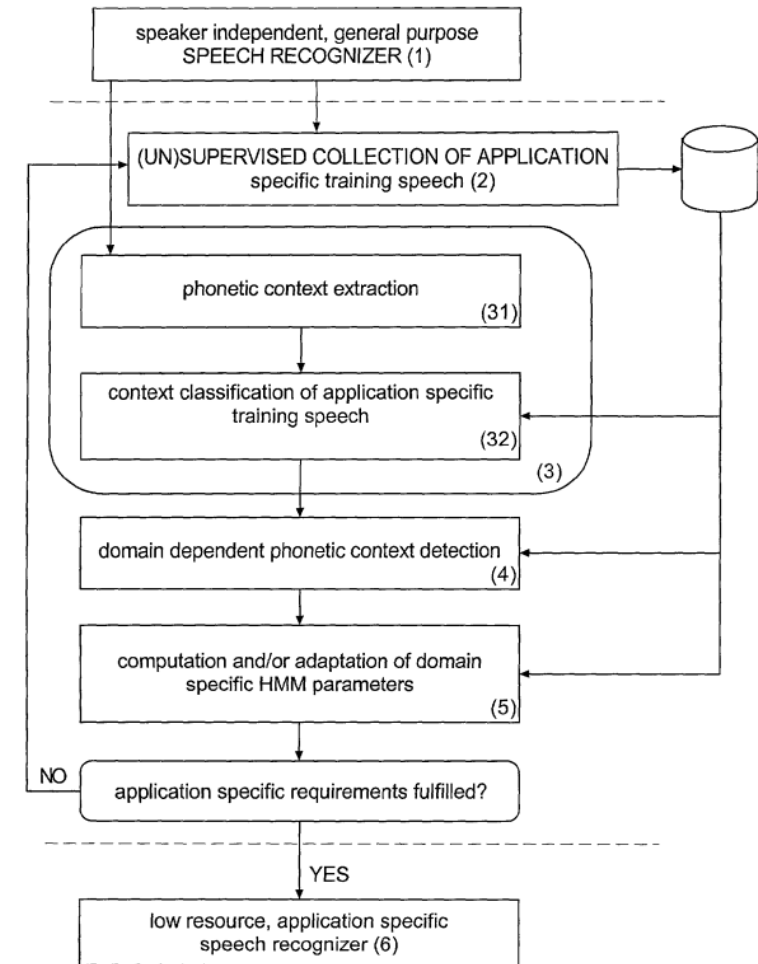


FIGURE 1

Specification Is Dispositive Here

“[T]he specification `is always highly relevant to the claim construction analysis. Usually, it is dispositive; it is the single best guide to the meaning of a disputed term.”“

Phillips v. AWH Corp., 415 F.3d 1303, 1315 (Fed. Cir. 2005)

Repeated descriptions of something as the present invention is definitional, as “[t]he public is entitled to take the patentee at his word”

Honeywell Intern., Inc. v. ITT Industries, Inc., 452 F. 3d 1312, 1318 (Fed. Cir. 2006)

Pacing Techs. LLC v. Garmin Int’l, Inc., 778 F.3d 1021, 1024-25 (Fed. Cir. 2015)

Rembrandt Patent Innovations, LLC v. Apple, Inc., 716 F. App’x 965, 972 (Fed. Cir. 2017)

“[W]here the specification makes clear that the invention does not include a particular feature, that feature is deemed to be outside the reach of the claims of the patent, even though the language of the claims, read without reference to the specification, might be considered broad enough to encompass the feature in question.”

Honeywell at 1320; *Pacing Techs.* at 1024-25

SciMed Life Sys. v. Advanced Cardiovascular Sys., 242 F.3d 1337, 1341 (Fed.Cir. 2001)

Prosecution History Also Confirms Generating by Re-estimation of the First Phonetic Contexts

- ① Applicants have also added new independent claim 28 and new dependent claims 30-32. **These claims emphasize the subject matter that is indicated as allowable in the Office Action.** Specifically, the new claims emphasize that the second speech recognizer generated from the first speech recognizer and the domain-specific data has a larger domain than the first speech recognizer.
- ② **These claims are supported by page 11, lines 8-15 and by previously claims 1, 13, 15 and 27.**
- ③ No new matter has been added as a result of these amendments. Feb. 3, 2005 Remarks, D.I. 84-5 at 11 (emphasis added)

- ① **Allowable subject matter required “all the limitations” of claim 13, including “by re-estimation of the first decision network and phonetic contexts” from claim 1:**

Claims 13 and 27 are objected to as being dependent upon a rejected base claim, but would be allowable **if rewritten in independent form including all of the limitations of the base claim and any intervening claims.**

Nov. 13, 2004 Office Action, D.I. 84-3 at 9-10 (emphasis added)

- ② **Support was exclusively from claims with generation by re-estimation**
- ③ **No new matter was added – meaning no other forms of generation**

Prosecution History Also Confirms Generating by Re-estimation of the First Phonetic Contexts

Applicant further describes the “invention” as the modification of acoustic or phonetic contexts from the first speech recognizer, i.e. re-estimation

Prior to addressing the rejection to the claims, a brief review of the Applicant’s claimed invention may be helpful. According to the Applicants’ claimed invention a first speech recognizer and domain-specific data can be combined to generate a second speech recognizer. The first speech recognizer can include a first decision network and its corresponding phonetic content. The second speech recognizer can include a second decision network and corresponding phonetic content. The structure and number of nodes within the second decision network can be different from the structure and number of nodes of the first decision network. More specifically, the Applicants’ invention can identify new acoustic contexts where needed, and can introduce an appropriate number of new context-dependent HMMs or HMM states (based on, for example, a split-and-merge procedure with domain or data specific thresholds).

Feb. 3, 2005 Remarks, D.I. 84-5 at 11 (emphasis added)

Case 1:19-cv-01438-PES Document 141 Filed 07/15/20 Page 43 of 52

Omilia's Construction Aligns All of the Independent Claims with the Invention Described in the Patent

PODS case cited by Nuance controlling. Claim 27 must have same meaning as the generation in claim 1. *PODS Inc. v. Porta Stor Inc.*, 484 F.3d 1359, 1366-67 (Fed. Cir. 2007)

Part of Claim 27 (at issue)

based on the first speech recognizer and the domain-specific training data, **generating a second acoustic model** of said first domain and said second domain comprising a second acoustic model with a second decision network and corresponding second phonetic contexts...

'925 patent at 14:15-20 (emphasis added)

Part of Claim 1 (representative of all other claims)

based on the first speech recognizer, **generating a second acoustic model** with a second acoustic model with a second decision network and corresponding second phonetic contexts for said second speech recognizer **by re-estimating said first decision network and said corresponding first phonetic contexts** based on domain-specific training data...

'925 patent at 10:48-53 (emphasis added)

“Automatically Generate”

Presentation by Omilia
No. Case No. 1:19-cv-11438-PBS
United States District Court
District of Massachusetts

Proposed Constructions

Claim Term (Purple) in Context	Asserted Claims	Omilia's Proposed Construction	Nuance's Proposed Construction
<p>“A computerized method of automatically generating from a first speech recognizer a second speech recognizer, . . . Said method comprising:”</p> <p>“[Machine-readable storage with a computer program] to automatically generate from a first speech recognizer a second speech recognizer, . . . Said machine-readable storage causing the machine to perform the steps of:”</p>	<p>925: 1, 2</p> <p>925: 14, 15</p>	<p>“generate/ing by a computer without human intervention”</p>	<p>“Generate/ing, at least in part by a computer”</p>

Case 1:19-cv-01148-PBS Document 1-1 Filed 07/15/20 Page 46 of 52

“Automatically Generating” Means Generating Without Human Intervention

A computerized method of **automatically generating** from a first speech recognizer a second speech recognizer, . . . Said method comprising:

'925 patent at 10:15-20 (emphasis added)

The dispute here is how much of the generation of the speech recognizer is performed by a computer vs. a human to be “automatic”

Omilia: Entire generation must be automatic (“without human intervention”) consistent with the claim language, common meaning of “automatically” and the specification

OR

Nuance: Partly generated by a computer (“at least in part”) which renders “automatically” meaningless

Case 1:19-cv-01148-PBS Document 1-1 Filed 07/15/20 Page 47 of 52

“Automatically Generating” Means Generating Without Human Intervention

Law is clear that Omilia’s construction is correct – “automatically generating**” means generating “**without human intervention**”:**

- 1. Context of claim makes clear that “automatically” means more than just a part is computerized, the entire method here is “without human intervention”**
- 2. “Automatically” typically understood and construed as “without human intervention” – even Nuance’s cases confirm this**
- 3. Specification is consistent; does not describe human intervention but a computerized generation**

Omilia's Construction Necessary To Give Meaning to the Claim

Claim 1 (representative of claim 14 as well)

A **computerized method** of **automatically generating** from **a first speech recognizer a second speech recognizer**, said first speech recognizer comprising a first acoustic model with a first decision network and corresponding first phonetic contexts, and said second speech recognizer being adapted to a specific domain, **said method comprising**:
 based on said first acoustic model, generating a second acoustic model with a second decision network and corresponding second phonetic contexts for said second speech recognizer by re-estimating said first decision network and said corresponding first phonetic contexts based on domain-specific training data, wherein said first decision network and said second decision network utilize a phonetic decision free to perform speech recognition operations, wherein the number of nodes in the second decision network is not fixed by the number of nodes in the first decision network, and wherein said re-estimating comprises partitioning said training data using said first decision network of said first speech recognizer.

'925 patent at 10:42-61 (emphasis added)

- ① “Automatically” in the limiting preamble – describes the entire generation method claimed
- ② Claims already require a computer; “automatically” is more, its what the computer does
- ③ Generation of a second speech recognizer is what is automatic but also lists certain required steps regarding the acoustic model

Law is Clear that “Automatically” Means “Without Human Intervention”

Case 1:19-cv-11438-PBS Document 141 Filed 07/15/20 Page 49 of 52

Source	Term	Explanation
<i>CollegeNet, Inc. v. ApplyYourself, Inc.</i> , 418 F.3d 1225 (Fed. Cir. 2005)	“automatically “inserting” and “storing”	Affirmed district court that “automatically” meant, “once initiated, the function is performed by a machine without the need for manually performing that function” or as described by the Federal Circuit (“without human intervention”) (at 1236)
<i>Whitserve, LLC v. Comput. Packages, Inc.</i> , 694 F.3d 10 (Fed. Cir. 2012)	“automatically” “querying,” “generating,” “transmitting” and “receiving”	No dispute of the District Court’s construction of “once initiated, is performed by a machine without the need for manually performing that process, that is without the need for human intervention.” (at 8-10)
<i>Rembrandt Patent Innovations, LLC v. Apple, Inc.</i> , 716 F. App’x 965 (Fed. Cir. 2017)	No explicit recitation of “automatic” in the claims	“Automated” should be included in the claims because of its description as the “present invention,” which it understood to mean “without human intervention.” Also found support in the prosecution history distinguishing prior art as requiring human intervention (at 971-72)
<i>Sky Techs. LLC v. IBM, Inc.</i> , 2005 U.S. Dist. LEXIS 47689 (E.D. Tex. Sept. 6, 2005)	“automated negotiations engine”	Determined “the ordinary meaning of ‘automated’ to mean without human intervention” (at *33)
<i>Phoenix Licensing, LLC v. Advance Am.</i> , U.S. Dist. LEXIS 147141 (E.D. Tex. Oct. 24, 2016)	“automatically”	Construed as “performed by a computer without input from a human” because of the use of “automatically” in the specification and statements during the prosecution distinguishing human input (at *127)

The Specification Describes the Steps of Generating Are Performed by a Computer – Not a Human

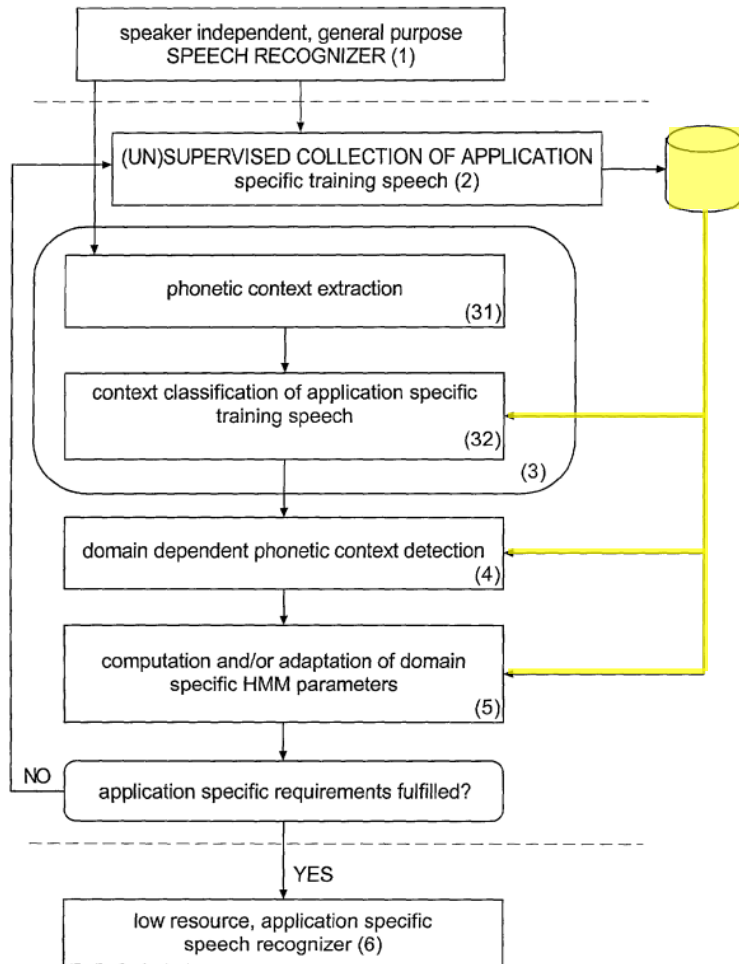


FIGURE 1

- **Abstract & Summary of invention use the same language, described as the “present invention” – “automatically generating” a second specialized speech recognizer from a first** Abstract, 2:31-34
- **Specification describes that the “present invention” makes use of a computer to “carry[] out the methods described.”** 3:9-21
- **Only description in the specification likewise shows the steps are performed by a computer** Fig. 1; Col. 7-8
 - No reference to a human performing steps of the generation
 - Collection of training data may or may not involve a human, however:
 - It occurs prior to generation, already collected for adaptation 7:27-36
 - Not within the scope of the claim. Claims just require the collected ed data (“based on”)
 - No support for suggestion that the setting of thresholds (7:66-8:3) and acoustic model parameter adaptation (element 4) done by a human
 - Just the opposite, spec shows a computer doing so
 - Also consistent with the dependent claims (5, 6): both steps of the “automatic” generation

Thank you

Dimitrios Drivas (admitted pro hac vice)
Raj Gandesha (admitted pro hac vice)
John Padro (admitted pro hac vice)
Stefan Mentzer (admitted pro hac vice)
WHITE & CASE LLP
1221 Avenue of the Americas
New York, NY 10020-1095
(212) 819-8286
ddrivas@whitecase.com
rgandesha@whitecase.com
john.padro@whitecase.com
smentzer@whitecase.com

Kevin C. Adam (SBN 684955)
Daniel S. Sternberg (SBN 688842)
WHITE & CASE LLP
75 State Street, 24th Floor
Boston, MA 02109
(617) 979-9300
kevin.adam@whitecase.com
daniel.sternberg@whitecase.com

Hallie Kiernan (admitted pro hac vice)
WHITE & CASE LLP
3000 El Camino Real
Two Palo Alto Square, Suite 900
Palo Alto, CA 94306
(650) 213-0300
hallie.kiernan@whitecase.com

Counsel for Omilia Natural Language
Solutions, Ltd.

